AFRL-AFOSR-JP-TR-2016-0014

Bio-Inspired Human-Level Machine Learning

Byoung-Tak Zhang
SEOUL NATIONAL UNIVERSITY
SNUR&DB FOUNDATION RESEARCH PARK CENTER
SEOUL, 151742
KR

12/21/2015
Final Report

Air Force Research Laboratory
Air Force Office of Scientific Research

Asian Office of Aerospace Research and Development
Unit 45002, APO AP 96338-5002

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
21-12-2015	Final	27/09-2012 to 26-09-2015

4. TITLE AND SUBTITLE
Bio-Inspired Human-Level Machine Learning

5a. CONTRACT NUMBER
FA2386-12-1-4087

5b. GRANT NUMBER
Grant 12RSZ110_124087

5c. PROGRAM ELEMENT NUMBER
61102F

6. AUTHOR(S)
Byoung-Tak Zhang

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
SEOUL NATIONAL UNIVERSITY
SNUR&DB FOUNDATION RESEARCH PARK CENTER
SEOUL, 151742
KR

8. PERFORMING ORGANIZATION REPORT NUMBER
N/A

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AOARD
UNIT 45002
APO AP 96338-5002

10. SPONSOR/MONITOR'S ACRONYM(S)
AFRL/AFOSR/IOA(AOARD)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
12RSZ110_124087

12. DISTRIBUTION/AVAILABILITY STATEMENT
A DISTRIBUTION UNLIMITED: PB Public Release

13. SUPPLEMENTARY NOTES

14. ABSTRACT
How can brain computation be so fast, flexible, and robust? What kinds of representational and organizational principles facilitate the biological brain to learn so efficiently and flexibly on the sub-second time scale and so reliably on the continuous lifetime scale? To understand these principles, we aimed to develop human-level machine learning technology that is fast, flexible, and reliable to adapt to a continuously changing, dynamic environment. Based on dynamic "neural"populations (neural assemblies), we constructed a "human-like" machine learning model and implement this model in "molecular" populations (molecular assemblies) using in vitro DNA computing. In the first year, we developed the dynamic hypernetwork models of neural populations in the sequential Bayesian framework for lifelong learning. In the second year, we extended it to the molecular dynamic hypernetwork model, and designed in vitro experimental protocols to implement online language learning from a stream of text corpus. In the third year, we demonstrated the use of molecular dynamic hypernetworks for multimodal visuo-linguistic concept learning from a long stream of video data and their extensions to high-level cognitive functions such as anagram solving problem.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Lt Col Brian Lutz
Unclassified	Unclassified	Unclassified	SAR	20	19b. TELEPHONE NUMBER *(Include area code)* +81-42-511-2000

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18

Final Report for AOARD Grant 124087

"Bio-Inspired Human-Level Machine Learning"

Oct/25/2015

Name of Principal Investigators (PI): Byoung-Tak Zhang
- E-mail address: btzhang@bi.snu.ac.kr
- Institution: Seoul National University
- Mailing Address: Room 302-323 (New Engineering Building), School of Computer Science and Engineering, Seoul National University, Seoul 151-744, Korea
- Phone: +82-2-880-1847
- Fax: +82-2-875-2240

Period of Performance: Sept/27/2012 – Sept/26/2015

Abstract:
How can brain computation be so fast, flexible, and robust? What kinds of representational and organizational principles facilitate the biological brain to learn so efficiently and flexibly on the sub-second time scale and so reliably on the continuous lifetime scale? To understand these principles, we aimed to develop human-level machine learning technology that is fast, flexible, and reliable to adapt to a continuously changing, dynamic environment. Based on dynamic "neural" populations (neural assemblies), we constructed a "human-like" machine learning model and implement this model in "molecular" populations (molecular assemblies) using *in vitro* DNA computing. In the first year, we developed the dynamic hypernetwork models of neural populations in the sequential Bayesian framework for lifelong learning. In the second year, we extended it to the molecular dynamic hypernetwork model, and designed in vitro experimental protocols to implement online language learning from a stream of text corpus. In the third year, we demonstrated the use of molecular dynamic hypernetworks for multimodal visuo-linguistic concept learning from a long stream of video data and their extensions to high-level cognitive functions such as anagram solving problem. We expect that the bio-inspired human-level machine learning combined with molecular-computing implementation can offer an interesting, novel paradigm to address for flexible and reliable computing.

Introduction:
One of the main challenges in artificial intelligence is to develop human-like machine learning technology that is fast, flexible, and reliable to adapt to a continuously changing, dynamic environment. Converging neuroanatomical and neurophysiological evidence shows that the brain uses distributed, overlapping representations based on sparse population codes that are coordinated dynamically (Averbeck et al., 2006; Pouget et al., 2000; von der Malsburg et al., 2010). We hypothesize that brain computation exploits the huge degrees of freedom generated by a large number of memory units, ranging from neurotransmitters and neurons to

cell-assembly, and organized into multiscale complex networks in space and coordinated dynamically in time (Caroni, 2012; Freeman, 2000).

The objective of this project is to build a learning-friendly computational model based on dynamic neural populations and implementing this model in self-assembling molecular populations using DNA computing. A key idea underlying this approach is that the plasticity of neural populations in the brain is based on molecular interactions at the physico-chemical level and, thus, molecular computational processes can naturally simulate human-like learning and memory. The molecular self-assembly mechanisms in DNA chemistry provide us a natural, physical medium for modeling dynamic "neural" populations (neural assemblies). Massively parallel mechanisms of *in vitro* DNA computing provide us a convenient tool for dealing with large populations, 10^{15} molecules in a nano-mole, which is bigger than the numbers of 10^{11} neurons and 10^{14} synaptic connections in the human brain.

In previous work, we experimentally demonstrated the feasibility of cognitive memory with DNA self-assembly. We showed that wet DNA computing can implement weighted-sum operations which are fundamental to perform pattern classification (Lim et al., 2010). Since pattern classification underlies many cognitive tasks, this work opened a new way of creating flexible cognitive memories *in vitro* with molecules. We also demonstrated the potential of the molecular self-assembly model to build associative language models automatically from language data to generate sentences (Lee et al., 2011).

On the mathematical and computational modeling side we developed a probabilistic graphical model of sparse, random population codes called hypernetworks (Zhang, 2008). The model also applied to a visually-grounded language learning (Zhang 35 al., 2012), where cognitive memory consists of multimodal compound concepts which are encoded as hyperedges (molecular memory particles) and then assembled, dissembled, and reassembled to be adapted incrementally as the video sequences are observed.

However, there were several challenges to achieving human-level learning and memory. First, the concept of population coding needed to be extended to deal with online, predictive learning in a changing environment. Second, representational formalisms and their translations between neural populations and molecular populations needed to be investigated. Third, the DNA computing and molecular learning technology needed be scaled up to make molecular computational simulation of the whole-brain scale, to make cognitive learning possible and to achieve human-level machine learning.

In the first year of the project, we focused on constructing mathematical theories of dynamic neural populations. Building upon our previous work on the hypernetwork models of cognitive learning and memory (Zhang, 2012), we developed population-coded dynamic hypernetwork models of lifelong learning in a non-stationary, changing environment [1, 2, 6, 8, 9, 17]. In [9], we discussed our model from the perspectives of embodied cognition, multisensory integration, cognitive dynamics, perception-action cycle, and lifelong learning. We developed a sequential Bayesian framework for lifelong learning, built a taxonomy of lifelong-learning paradigms, and examined information-theoretic objective functions for each paradigm, with an emphasis on active learning. Also, in [7], we presented that DNA hybridization can be modeled as computing the inner product between

embedded vectors in a corresponding vector space, and proposed the algorithm performing learning of a binary classifier in this vector space.

In the second year, we extended this to the molecular dynamic hypernetwork model, and designed *in vitro* experimental protocols to implement online language learning from a stream of text corpus [3, 4, 10, 14, 19, 20, 23]. To measure the difference between different information-encoded sequences, we introduced the symmetric internal loops of double stranded DNA, and which were used to recognize similar or different patterns. Through a series of training processes which is simply storing the given training data in different microtubes in each class of hypernetwork, we observed that the accuracy of sentence classification tasks increased on the corpus of TV show dialogue and our molecular learning was able to generalize the training sentences.

In the third year, we demonstrated the use of molecular dynamic hypernetworks for multimodal visuo-linguistic concept learning from a long stream of video data. Motivated by the cognitive developmental process of children constructing the visually grounded concepts from multimodal stimuli (Meltzoff, 1990), we proposed a hierarchical model of automatically constructing visual-linguistic knowledge by dynamically learning concepts represented with vision and language from videos [8, 12, 15, 16, 22]. We developed a stochastic method for graph construction, i.e. a graph Monte Carlo algorithm, and our model learns the concepts by the algorithm while observing new videos, thus robustly tracing concept drift and continuously accumulating new conceptual knowledge. Using a series of approximately 200 episodes of educational cartoon videos we examined the emergence and evolution of the concept hierarchies as the video stories unfold. Through the experiment, we observed that the number of visual and linguistic nodes tends to increase, because the concepts continuously develop while observing the videos. Also, we presented a molecular computational model for human anagram solving to show the potential of application to high-level cognitive functions [5, 11, 13, 18, 21].

Our major contribution is to propose the molecular assembly model of cognitive memory and learning which can be used as a tool for simulating cognitive dynamics involved with multisensory cue integration, grounded concept learning, and interaction of vision and language. We believe that the bio-inspired human-level machine learning combined with molecular-computing implementation can offer an interesting, novel paradigm to address for flexible and reliable computing. We also expect that the cognitive memory architectures and their learning algorithms contribute to revolutionize the AI technology to be used in lifelong learning, self-organizing, sensorimotor systems.

[1st Year] The Dynamic Hypernetwork Models of Neural Populations

Experiments:

In the first year, we constructed a dynamic Bayesian inference framework and examined information-theoretic objective functions for lifelong learning [9]. In lifelong learning, training data are observed sequentially as learning unfolds and not kept for iterative reuse. The learning is proceeded in an online and incremental manner over an extended period in a changing environment. This requires incremental transfer of knowledge acquired from previous learning to future learning, which can be formulated as a Bayesian inference. We applied a sequential Bayesian framework for lifelong learning to build taxonomy of lifelong-learning paradigms, and examine information-theoretic objective functions for each paradigm (Figure 1).

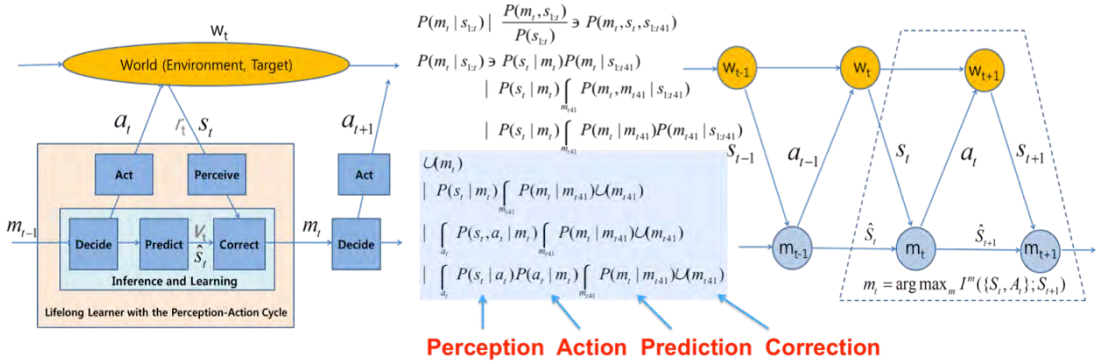


Figure 1. Lifelong learning with action-perception-learning cycle [9]

Results and Discussion:

We distinguished three paradigms of lifelong-learning: learning with passive and continual observations, learning with actions (but without reward feedbacks), and active learning with explicit rewards. For each of the paradigm we examined the objective functions of the lifelong learning styles: prediction errors and predictive information, empowerment which measures how much influence an agent has on its environment, and the value function or the expected reward of the agent.

We believe the general framework and the objective functions for lifelong learning can provide a baseline for evaluating the representations and strategies of the learning algorithms. Specifically, the objective functions can be used for innovating algorithms for discovery, revision, and transfer of knowledge of the lifelong learners over the extended period of experience. Our emphasis on information theory-based active and predictive learning with minimal mechanistic assumptions on model structures can be especially fruitful for automated knowledge acquisition and sequential knowledge transfer between a wide range of similar but significantly different tasks and domains.

For a theoretical study, we also presented a computational learning method for bio-molecular classification [7]. In this study, we showed how to design biochemical operations both for learning and pattern classification. DNA hybridization is modeled as computing the inner product between embedded vectors in a corresponding vector space (Figure 2), and our algorithm performed learning of a binary classifier in this vector space. Our algorithm manipulates populations of DNA sequences via hybridization and denaturing operations, modifying distributions of the associated

vectors in the kernel feature space. After learning is performed on data examples, an unknown DNA sequence molecule can be directly classified using the learned weights in the molecular population. We analyzed the thermodynamic behavior of these learning algorithms, and showed simulations on artificial and real datasets as well as demonstrate preliminary wet experimental results using gel electrophoresis.

In our classification results with the generated data shown in Figure 3, points in a two-dimensional space are labeled into two classes shown in the yellow and blue color. In this space, the binding energy is given by the Euclidean distance between pairs of points. The contours represent various hybridization amounts, and change according to the annealing temperature schedules. This shows how controlling the hybridization schedule influences both the positive definiteness and sparsity of the resulting kernel matrices. With sufficient annealing as shown in Figure 3(a), the kernel satisfies positive definiteness. In Figure 3(c) with no annealing, the kernel does not satisfy positive definiteness, resulting in bad classification results. With high temperature hybridization in Figure 3(b), the kernel matrix is positive definite but very diagonally dominant and sparse. In this case, the hybridization contours show that the decision surface depends more specifically on nearest neighbors as compared to the decision surface in Figure 3(a). Such a sparse kernel matrix would be more vulnerable to noise in the training data. The ROC curves showed that the classification performance of our proposed method is superior to kFDA and performs better than the SVM algorithm (Figure 4).

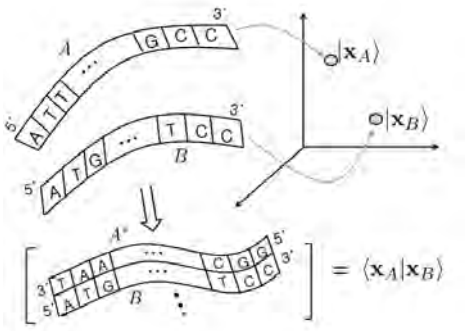


Figure 2. DNA sequence mapped into a vector space by an inner product [7]

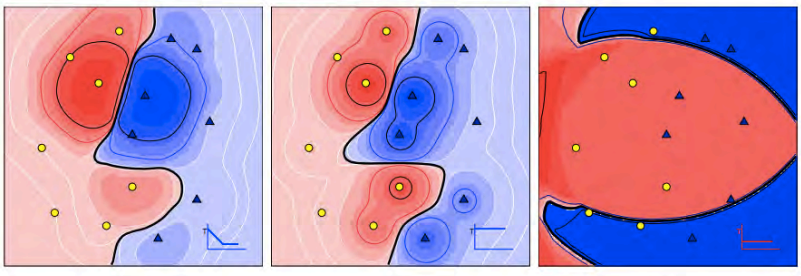


Figure 3. Classification of two-class data learned with different temperature schedules [7] (a) 80°C to 20°C, (b) 80°C constant, and (c) 30°C constant

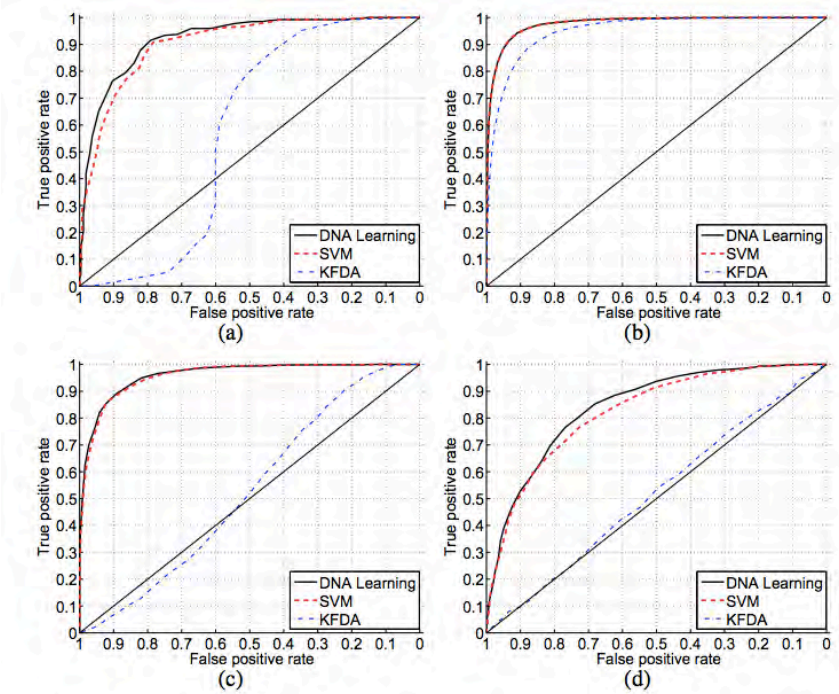


Figure 4. The classification results performed using DNA learning, SVMs and kernel FDA using the same DNA kernel [7]

[2nd year] DNA-Computing Implementations of the Dynamic Hypernetwork Models

Experiments:

In the second year, we developed a molecular machine learning model *in vitro* using symmetric internal loops of double stranded DNA [23]. To enable the molecules to learn, the way of measuring differences between sequences was needed. By using mismatching DNA sequences during hybridization, we encoded information into DNA sequences (Figure 5) and designed the DNA sequences to produce the symmetric internal loops when matched with the different sequences of same size. (Figure 6) These mismatches were used to determine the distances between given instances, which is essential for recognizing similar or different patterns.

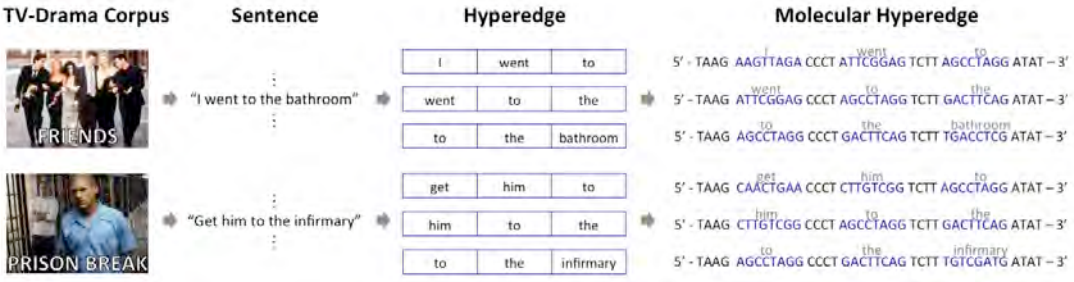


Figure 5. Encoding sentences in DNA sequences in the structure of hyperedge

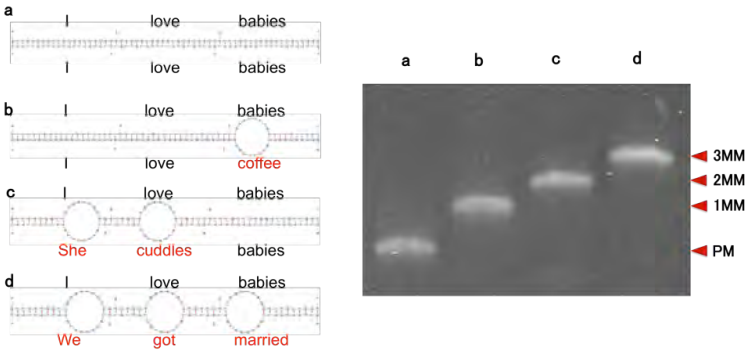


Figure 6. Symmetric internal loops of double stranded DNA in sentences [23(submitted)]

The training process involves simply storing the given training data in different microtubes in each class of hypernetwork. When a new training data is encountered, similar and identical instances are retrieved from the hypernetworks and used to classify the new example. The classification of the data is conducted through gel electrophoresis by comparing relative intensity of the bands (Figure 7). The intensity of the band represented the probability of that test data to be classified into that class, i.e. a higher band intensity meant higher probability that the sentence belonged to the according class.

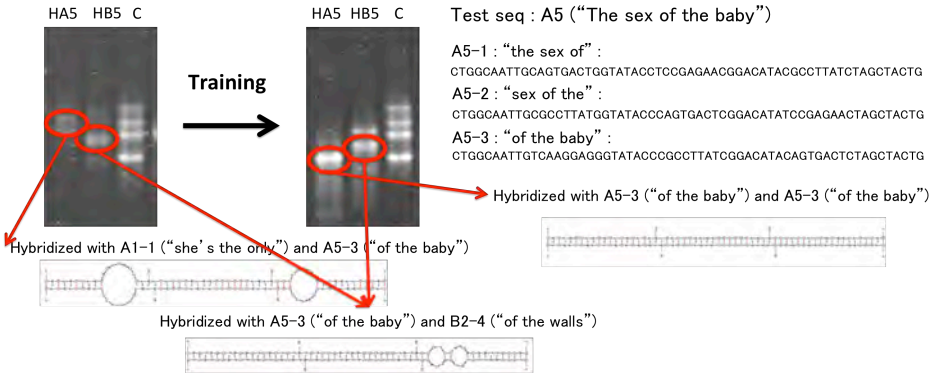


Figure 7. Classification after training [23(submitted)]

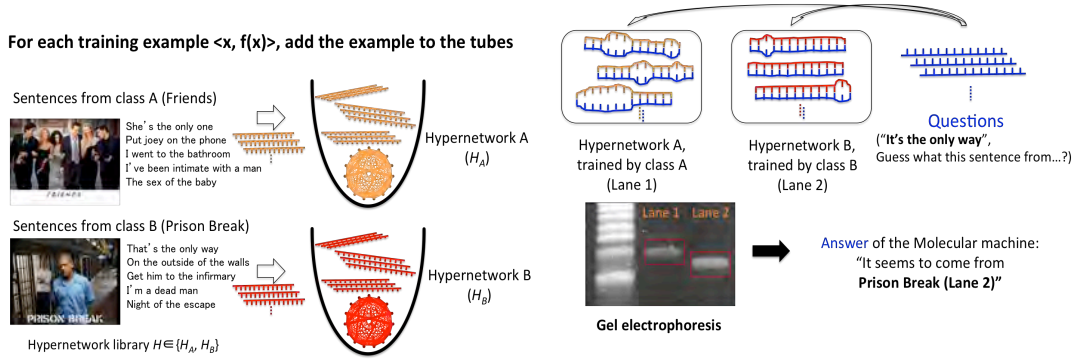


Figure 8. The proposed molecular machine learning process [23(submitted)]

To evaluate the model, DNA molecules were trained using a set of sentences obtained from a corpus of TV drama dialogue and tested using a set of unknown sentences from same corpus (Figure 8). We collected sentences of the TV drama videos of 'Friends' and 'Prison Break' for learning and testing the DNA hypernetwork models. We designed the DNA sequences for implementing the

molecular hypernetwork model of language that can distinguish the sentences whether they come from Friends or Prison Break. A 20-sentence classification experiment has been conducted to evaluate the feasibility of the population-coding based molecular learning of language concepts. 10 of the 20 sentences (5 from Friends, 5 from Prison Break) were used to train, and the other 10 sentences (5 from Friends, 5 from Prison Break) were used to test.

Results and Discussion:

The result of our experiments showed that the molecular learning machine was able to generalize training sentences (Figure 9). We summed up the correctly classified examples in each classification test and presented these results in bar graphs (Figure 10). The hypernetwork was gradually trained and tested at each step. Regardless of the low accuracy in the initial training phase, the accuracy was increased to 100% at the end of the training process in each case.

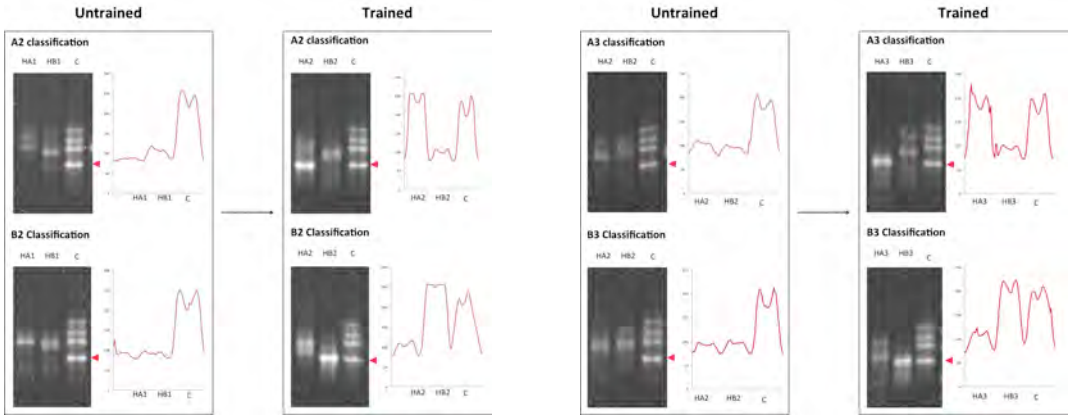


Figure 9. Verification of training steps by classification of (A) Friends and (B) Prison Break training data [23(submitted)]

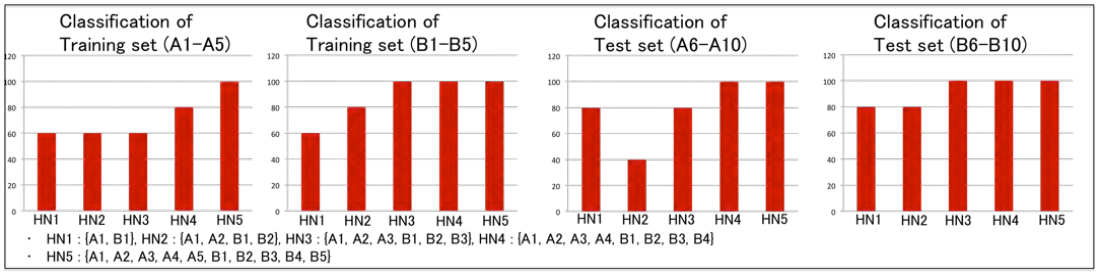


Figure 10. Accuracy of the classification of test and training examples in each training step [23(submitted)]

The major contribution of this work is the implementation of machine learning algorithm *in vitro* exploiting the symmetric internal loops. We verified each molecular learning step and performed classification experiments using the test data, which enabled to present the generalization phenomenon. By exploiting the generality of machine learning, our novel molecular learning machine could in principle be used to solve other problems such as text mining and molecular recognition in biology if the data can be properly encoded in DNA molecules.

[3rd Year-1] Molecular Dynamic Hypernetworks for Multimodal Concept Learning

Experiments:

In the third year, we applied the molecular dynamic hypernetwork models to learning multimodal vision-language concepts from videos. The resulting model is called deep concept hierarchy (DCH) [16] and consists of two or more concept layers and one layer of multiple modalities (Figure 11). Each concept layer is represented by a hypergraph structure, and this structure enables the multiple levels of concepts to be represented by the probability distribution of the visual-textual variables (Figure 12). The higher concept layers represent more abstract concepts than the lower layers, and the modality layer contains the populations of many microcodes encoding the higher-order relationships among two or more visual and textual variables. Each concept is represented as the probability distribution of word-patch appearance.

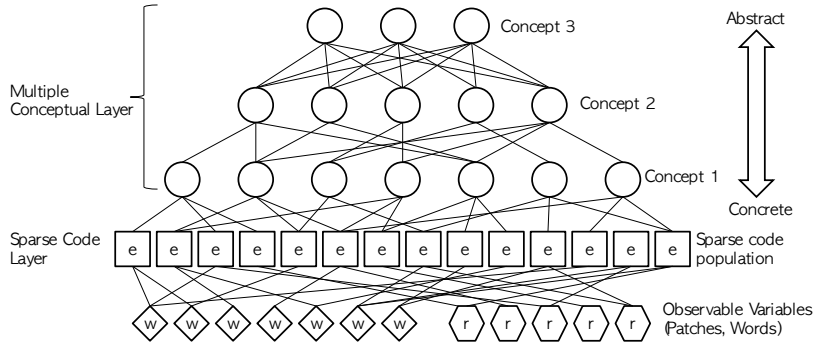


Figure 11. Architecture of deep concept hierarchy [16]

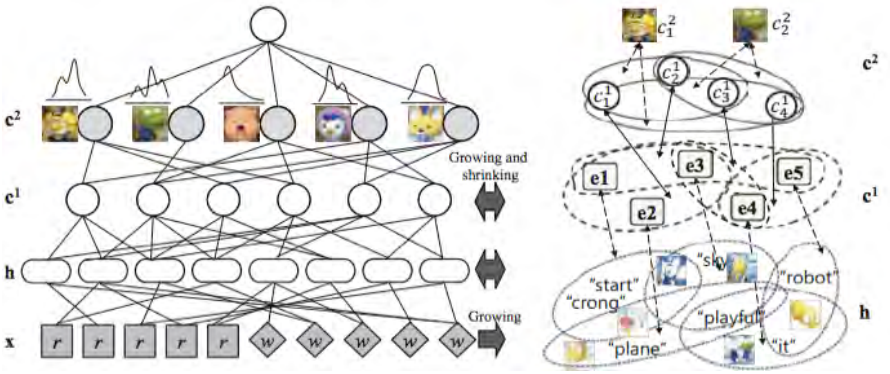


Figure 12. Example of deep concept hierarchy learned from Pororo videos [16]

To efficiently search the huge space of the vision-language concepts, we developed a stochastic method for graph construction, i.e. a graph Monte Carlo (graph MC) algorithm. DCH incrementally learns the concepts by the graph MC and the weight update process while observing new videos, thus robustly tracing concept drift and continuously accumulating new conceptual knowledge, allowing for being deployed in lifelong learning environments.

To verify our proposed model, the experiments conducted using the collection of the cartoon video "Pororo" consisting of 183 episodes with 1,232 minutes of playing time. As training and test data, 16,000 picture-sentence pairs were prepared, and we ran cognitive developmental experiments using population-coded hypernetwork

models to see how concepts are formed and revised incrementally as more episodes of video are watched sequentially. We visualized the developmental process of concepts in hypernetworks, such as 'train' and 'rabbit' as they emerge, disappear, and reemerge as the topic changes over the long period of sequential learning process.

	1~13 episodes (1 DVD)			1~183 episodes (14 DVDs)		
Concepts	Visual nodes	# of nodes (V/L)	Top 15 linguistic nodes	Visual nodes	# of nodes (V/L)	Top 15 linguistic nodes
Pororo		986/230	crong, you, clean, over, draw, huh, to, it, I, up, said, the, moving, is, pororo		12870/1031	crong, you, snowboarding, transforming, rescuing, pororo, the, lamp, seven, are, quack, yellow, not, lollipop, cake,
Eddy		644/198	I, ear, art, midget, game, nothing, say, early, diving, lost, middle, lesson, case, because, snowballs		9008/860	transforming, I, hand, careful, throw, art, suit, midget, farted, reverse, stage, luggage, gorilla, pole, cannon
Tongtong	-	0/0	-		1812/429	kurikuri, doodle, doo, avoid, airplane, crystal, puts, branch, bland, finding, pine, circle, kurikuritongtong, bees, talent

Figure 13. Visual-linguistic representation and development of 3 character concepts of video contents [16]

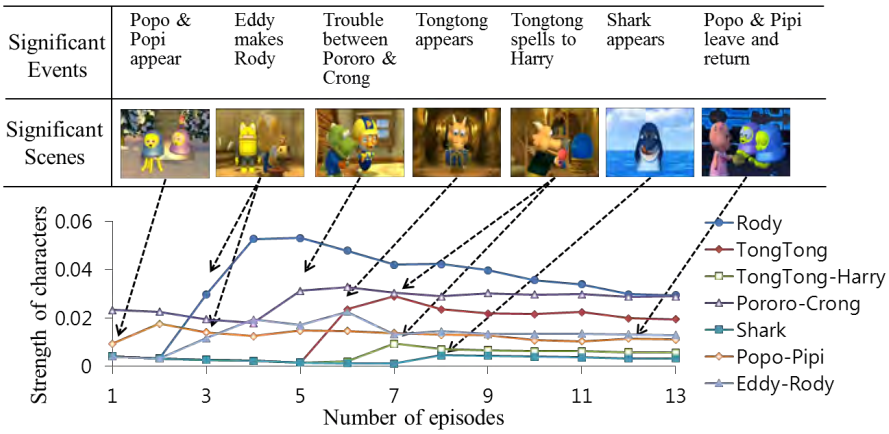


Figure 14. Change of strength of characters as the number of the observed episodes are increased [16]

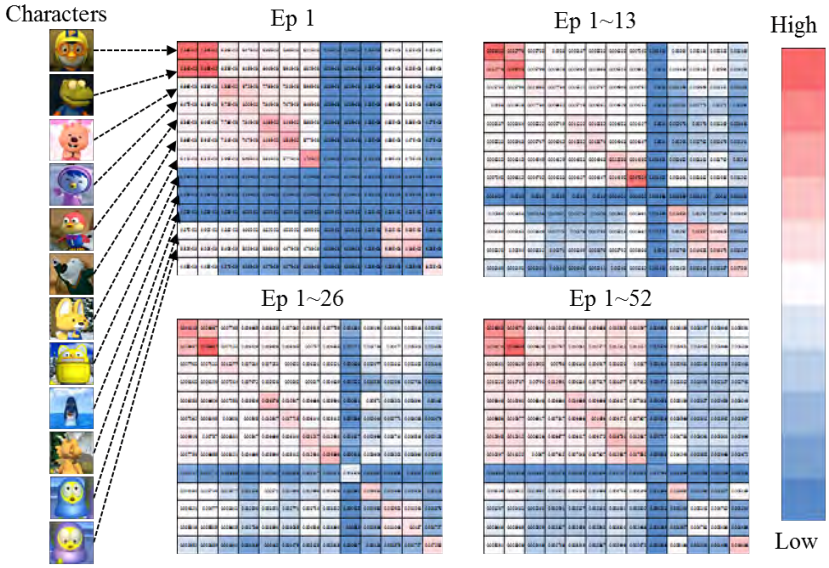


Figure 15. Change of the character relationships as all the stories unfold [Ha et al., 2015].

Results and Discussion:

Using a series of approximately 200 episodes of educational cartoon videos we examined the emergence and evolution of the concept hierarchies as the video stories unfold. Through the experiment, we observed that the number of visual and linguistic nodes tends to increase (Figure 13), because the concepts continuously develop while observing the videos (Figure 14, Figure 15).

We also presented the application of the deep concept hierarchies for context-dependent translation between vision and language, i.e. the transcription of a visual scene into text and the generation of visual imagery from text. In the scene-to-sentence generation experiment (Figure 16), we observed that the different subtitle was retrieved when query is given with and without character information. This means that our proposed model considers character information from visual image through concept learning. In the experiment of scene image generation from given query sentences (Figure 17), the generated scenes were synthesized by the weighted overlapping of image patches associated with the words in the sentences based on the constructed knowledge. As the number of observed videos increase, the images become more complex and diverse.

Subtitle	TongTong Let me take it to clock
Query	 translated query = {me, take}
With char info	- **tong tong tong let me take it to normal quack quack** - **tong tong let me take it i'm out of here come inside a cave in the forest**
Without char info	- don't mind him i'm sure it'll make you were lying to me take it loopy - huh no you don't like listening to the last time he let me take a look at poby listen

Figure 16. Example of scene-to-sentence generation result [16]

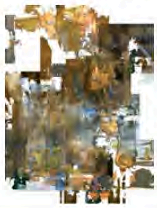
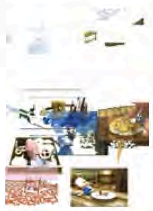
Query sentences	**1~52 episodes (1 season)**	**1~104 episodes (2 seasons)**	**1~183 episodes (all seasons)**
• Tongtong, please change this book using magic. • Kurikuri, Kurikuri-tongtong!			
• I like cookies. • It looks delicious • Thank you, loopy			

Figure 17. Example of sentence-to-scene generation result [16]

[3rd Year-2] Extension to a High-Level Cognitive Function: Anagram Solving Problem

Experiments:

We also applied the molecular dynamic hypernetwork model to solving the anagram solving. An anagram is a word play to find a new word from given alphabet letters. Good anagram solvers use the strategy of a solving parallel constraint satisfaction using bigrams, whereas poor anagram solvers use the strategy of serial searching all the possible arrangements of the given alphabets (Figure 18).

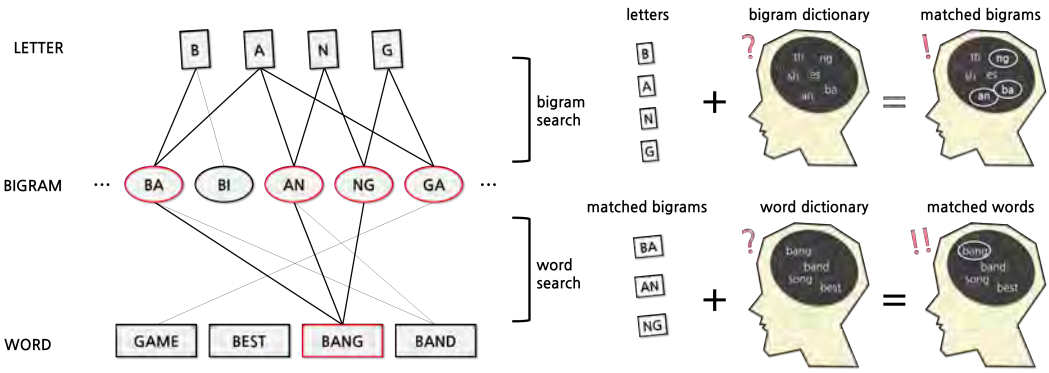


Figure 18. Good anagram solver's solving strategy

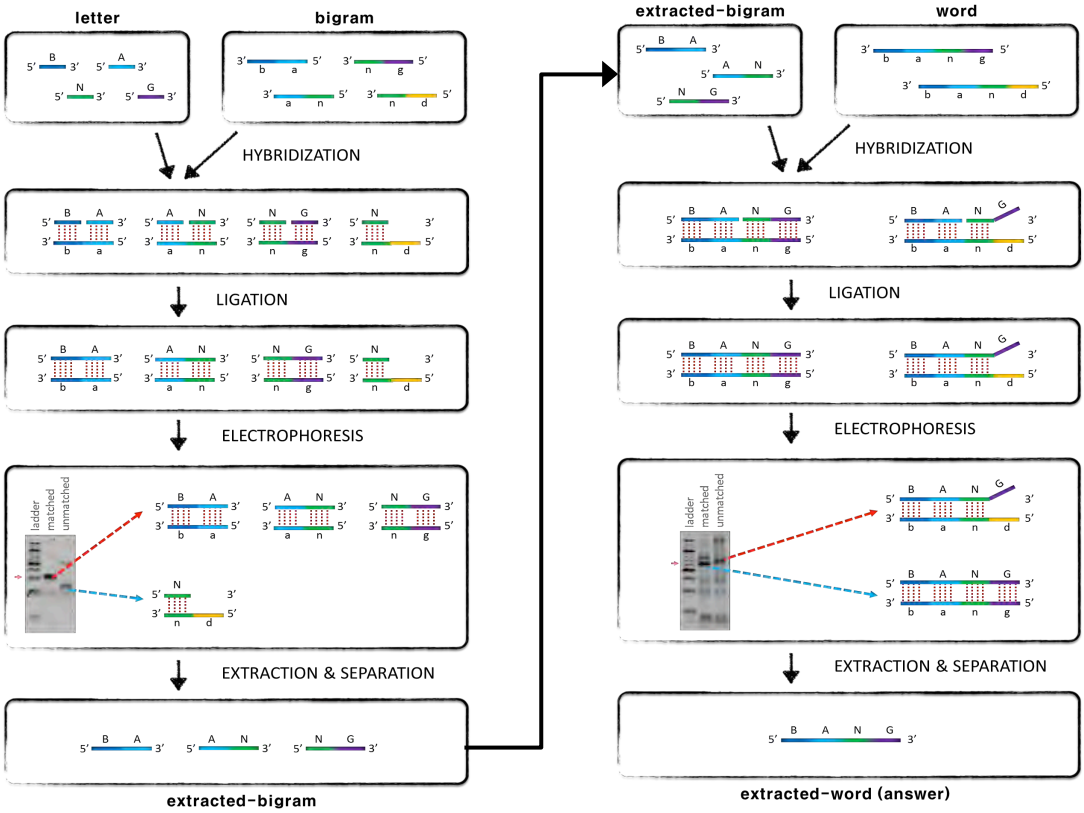


Figure 19. Molecular anagram solving algorithm

To demonstrate the capability of solving high-level cognitive function, we proposed a molecular computational algorithm of the good solver's solving process (Figure 19)

[5, 11]. We encoded letters into DNA sequences and made bigrams and words connecting the letter sequences. From letters and bigrams, we performed DNA hybridization. Each alphabet strand binds to its complementary bigram strand in parallel during this process. Then, ligation, gel electrophoresis, extraction and separation to extract matched bigrams. From the matched bigrams and words, we performed the above molecular operations again to distinguish between the right and wrong word.

To evaluate our model, we conducted a computational simulation and wet-lab experiment. In the computational simulation, we used the TV drama 'Friends' corpus to construct the bigram dictionary and word dictionary, and compared the speed of finding answers when using bigrams (i.e., the strategy of good solvers) and not using bigrams (i.e., the strategy of poor solvers). In the wet-lab experiment, we gave the set of words to the anagram solver, and tested if it can tell the difference between the correct answers and wrong answers.

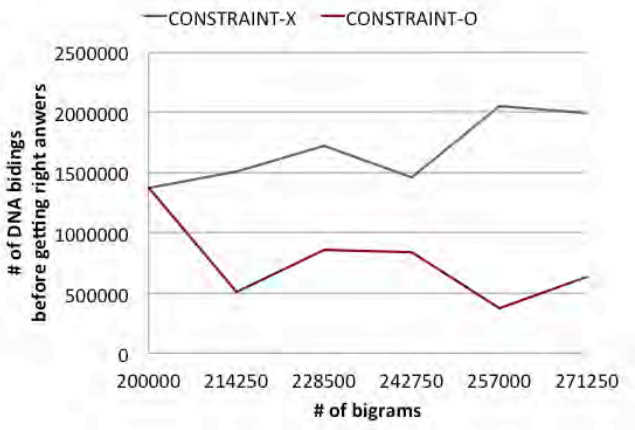


Figure 20. Computational simulation results of anagram solving [11]

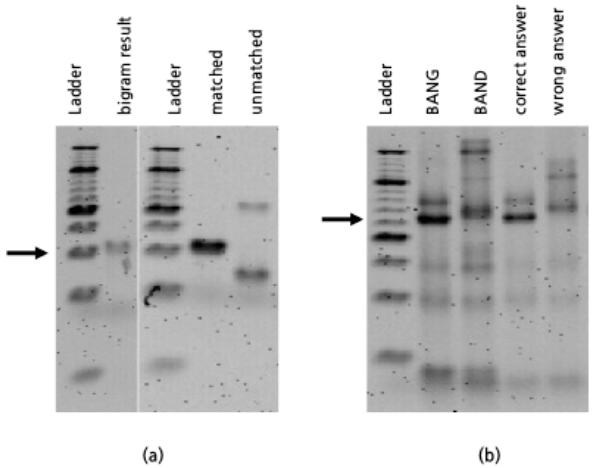


Figure 21. Molecular simulation results of anagram solving [5]

Results and Discussion:

Through the computational simulation [11], the experimental result showed that the good anagram solvers tend to come up with solutions faster than poor solvers who are likely to perform a serial hypothesis-testing process for a solution (Figure 20).

In the wet-lab experiment [5], our molecular anagram solver could tell the difference between the correct answers and wrong answers (Figure 21). The molecular anagram solver showed a higher intensity of gel-bands in the lane of the correct answers than the lane for the wrong answers, after conducting the series of molecular operations, including DNA hybridization, ligation, gel electrophoresis, extraction and separation.

This work proposed a new application for molecular computing that simulates the cognitive and parallel thinking process of humans and opens up the possibility for being used as a useful tool for computational modeling of cognitive processes.

References

Averbeck, B.B., Latham, P.E., & Pouget, A., Neural correlations, population coding and computation, *Nature Reviews Neuroscience*, 7:358-366, 2006.

Caroni, P., Donato, F., & Muller, D., Structural plasticity upon learning: Regulation and functions, *Nature Reviews Neuroscience*, 13:478-490, 2012.

Freeman, W. J., *Neurodynamics: An Exploration in Mesoscopic Brain Dynamics*, Springer, 2000.

Lee, J.-H., Lee, S.H., Chung, W.-H., Lee, E.-S., Park, T.H., Deaton, R., & Zhang, B.-T., A DNA assembly model of sentence generation, *Biosystems*,106: 51-56, 2011.

Lim, H.-W., Lee, S.H., Yang, K.-A., Lee, J.Y., Yoo, S.-I., Park, T.H., & Zhang, B.-T., *In vitro* molecular pattern classification via DNA-based weighted sum operation, *Biosystems*, 100(1):1-7, 2010.

Meltzoff, A. N., Toward a Developmental Cognitive Science: The Implications of Cross-modal Matching and Imitation for Development of Representation and Memory in Infancy, *Annual New York Academy Science*, 608:1-31, 1990.

Pouget, A., Dayan, P. & Zemel, R., Information processing with population codes, *Nature Reviews Neuroscience*, 1:125–132, 2000.

von der Malsburg, C., Phillips, W.A., & Singer, W. (Eds.), *Dynamic Coordination in the Brain: From Neurons to Mind*, MIT Press, 2010.

Zhang, B.-T., Hypernetworks: A molecular evolutionary architecture for cognitive learning and memory, *IEEE Computational Intelligence Magazine*, 3(3):49-63, 2008.

Zhang, B.-T., Ha, J.-W., & Kang, M.-G., Sparse population code models of word learning in concept drift, In *Proceedings of the 34th Annual Meeting of the Cognitive Science Society* (CogSci 2012), 2012.

List of Publications and Significant Collaborations that resulted from AOARD supported project:

a) Papers published in peer-reviewed journals

[1] Kim, S.-J., Ha, J.-W., Zhang, B.-T., Constructing higher-order miRNA-mRNA interaction networks in prostate cancer via hypergraph-based learning, *BMC Systems Biology*, 7:47, 2013.

[2] Shin, S.-Y., Yang, K.-A., Lee, I.-H., Lee, S.H., Park, T.H., Zhang, B.-T., Rule-based in vitro molecular classification and visualization, *BioChip Journal*, 7(1), 2013.

[3] Lee, I.-H., Lee, S.H., Park, T.H., Zhang, B.-T., Non-linear molecular pattern classification using molecular beacons with multiple targets, *BioSystems*, 114(3):206-213, 2013.

[4] Lim, H.-W., Lee, S.H., Yang, K.-A., Yoo, S-I., Park, T.H., Zhang, B.-T., Biomolecular computation with molecular beacons for quantitative analysis of target nucleic acids, *BioSystems*, 111(1):11-17, 2013.

[5] Chun, H.-S., Lee, J.-H., Ryu, J.-H., Baek, C., Zhang, B.-T., Molecular computing simulation of cognitive anagram solving, *Journal of the Korean Institute of Information Scientists and Engineers: Computing Practices*, 20(12):700-705, 2014(in Korean).

[6] Kim, S.-J., Ha, J.-W., Zhang B.-T., Bayesian evolutionary hypergraph learning for predicting cancer clinical outcomes, *Journal 6 Biomedical Informatics*, 49:101-111, 2014.

[7] Noh, Y.-K., Lee, D.D., Yang, K.-A., Kim, C., Zhang, B.-T., Molecular learning with DNA kernel machines, *BioSystems*, 2015(in press).

b) Papers published in peer-reviewed conference proceedings

[8] Lee, B.-J., Ha, J.-W., Kim, K.-M., Zhang, B.-T., Evolutionary concept learning from cartoon videos by multimodal hypernetworks, *In Proceedings of the IEEE Congress on Evolutionary Computation (CEC 2013)*, 1186-1192, 2013.

[9] Zhang, B.-T., Information-theoretic objective functions for lifelong learning, *In Proceedings of the AAAI 2013 Spring Symposium on Lifelong Machine Learning*, 62-69, 2013.

[10] Ryu, J.-H., Lee, J.-H., Zhang, B.-T., Integrated DNA Encoding of Semantic and Orthographic Distances between words, *In Proceedings of the 40th KIISE Fall Conference*, 683-685, 2013(in Korean).

[11] Lee, J.-H., Chun, H.-S., Lee, E. S., Ryu, J.-H., Zhang, B.-T., Simulation of Anagram Solving by Molecular Computer, *In Proceedings of the 40th KIISE Fall Conference*, 723-725, 2013(in Korean).

[12] Kim, K.-M., Ha, J.-W., Lee, B.-J., Zhang, B.-T., Character-based subtitle generation by learning of multimodal, *In Proceedings of the KIISE Korea Computer Congress (KCC2014)*, 647-649, 2014(in Korean).

[13] Chun, H.-S., Lee, J.-H., Ryu, J.-H., Baek, C., Zhang, B.-T., Molecular experimental simulation for cognitive anagram solving, *In Proceedings of the KIISE Korea Computer Congress (KCC2014)*, 849-850, 2014(in Korean).

[14] Lee, J.-H., Chun, H.-S., Zhang, B.-T., Simulation for molecular pattern classification of drama sentences based on DNA molecules, *In Proceedings of the KIISE Korea Computer Congress (KCC2014)*, 883-884, 2014(in Korean).

[15] Zhang, B.-T., Ontogenesis of agency in machines: a multidisciplinary review, *AAAI Fall Symposium*, 2014.

[16] Ha, J.-W., Kim, K.-M., Zhang, B.-T., Automated construction of visual-linguistic knowledge via concept learning from cartoon videos, *In Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence (AAAI 2015)*, 522-528, 2015.

c) Papers published in non-peer-reviewed journals and conference proceedings

d) Conference presentations without papers

[17] Zhang, B.-T., Higher-order predictive information for learning an infinite stream of episodes, *NIPS-2012 Workshop on Information in Perception and Action, Neural Information Processing Systems (NIPS)*, Lake Tahoe, 2012.

[18] Lee, J.-H., Lee, E.S., Ryu, J.-H., Chun, H.-S., Zhang, B.-T., Molecular computational simulation of cognitive processes for anagram solving, *The 19th International Conference on DNA Computing and Molecular Programming (DNA 19)*, Arizona, 2013.

[19] Ryu, J.-H., Lee, J.-H., Zhang, B.-T., Integrated encoding of semantic and orthographic distances in a DNA language model, *The 19th International Conference on DNA Computing and Molecular Programming (DNA 19)*, Arizona, 2013.

[20] Ryu, J.-H., Lee, J.-H., Zhang, B.-T., Development of molecular rewrite operation using Mung Bean enzyme, *The 20th International Conference on DNA Computing and Molecular Programming (DNA 20)*, Kyoto, 2014.

[21] Chun, H.-S., Lee, J.-H., Ryu, J.-H., Baek, C., Zhang, B.-T., Molecular computational simulation for human anagram solving process, *The 20th International Conference on DNA Computing and Molecular Programming (DNA20)*, Kyoto, 2014.

[22] Chun, H.-S., Zhang, B.-T., DNA encodings for in vitro molecular learning of multimodal data, *The 21st International Conference on DNA Computing and Molecular Programming (DNA 21)*, Cambridge, 2015.

e) Manuscripts submitted but not yet published

[23] Lee, J.-H., Baek, C., Chun, H.-S., Ryu, J.-H., Kim, J.-W., Deaton, R., Lee, S. H., Park, T. H., Zhang, B.-T., In vitro molecular machine learning via symmetric internal loops of DNA, *PLOS ONE,* 2015(submitted).

f) Provide a list any interactions with industry or with Air Force Research Laboratory scientists or significant collaborations that resulted from this work.

Attachments: Publications a), b) and c) listed above.

DD882: *As a separate document, the inventions disclosure form is completed and signed.*